

Listing of Claims:

1. (Currently Amended) A computer ~~assisted~~/implemented method for developing a classifier for classifying electronic communications comprising:

querying a user for an irrelevancy keyword indicative of an irrelevant electronic communication;

receiving a user identification of the irrelevancy keyword;

(a) presenting a user-generated electronic communications to a user for labeling as relevant or irrelevant, the electronic communications being selected from groups of electronic communications including:

a training set group of electronic communications, the training set group of electronic communications being selected by an active learning algorithm;

a system-labeled set of electronic communications previously labeled by the system;

a test set group of electronic communications, the test set group of electronic communications for testing the accuracy of a current state of a classifier being developed;

a faulty set of electronic communications suspected to be previously mis-labeled by the user; and

a random set of electronic communications previously labeled by the user;

(b) developing the classifier for classifying electronic communications based upon the irrelevancy keyword, the ~~relevant/irrelevant~~relevant labels and the irrelevant labels assigned by the user during the presenting of the electronic communications to the user; ~~and~~

(c) deploying the classifier for use in classifying electronic communications based upon the ~~relevant/irrelevant~~relevant labels and the irrelevant labels; and

(d) storing a set of electronic communications labeled by the classifier in a memory.

2. (Currently Amended) The method of claim 1, wherein the presenting of the electronic communications to the user includes:

assessing a value related to performance that labeling a set of electronic communications from each group provides to the classifier being developed; and

selecting a next group for labeling based upon a greatest respective value provided to the classifier being developed from the assessing.

3. (Currently Amended) A computer ~~assisted~~/implemented method for developing a classifier for classifying electronic communications comprising:

querying a user for an irrelevancy keyword indicative of an irrelevant electronic communication;

receiving a user identification of the irrelevancy keyword;

(a) presenting electronic communications to a user for labeling as relevant or irrelevant, the electronic communications being selected from groups of user-generated electronic communications including:

a training set group of electronic communications, the training set group of electronic communications being selected by an active learning algorithm;

a test set group of electronic communications, the test set group of electronic communications for testing the accuracy of a current state of a classifier being developed; and

a previously-labeled set of electronic communications previously labeled by at least one of the user, the system and another user;

(b) developing the classifier for classifying electronic communications based upon the irrelevancy keyword and the ~~relevant/irrelevant~~relevant labels and the irrelevant labels assigned by the user; ~~and~~

(c) deploying the classifier for use in classifying electronic communications based upon the ~~relevant/irrelevant~~relevant labels and the irrelevant labels; and

(d) storing a set of electronic communications labeled by the classifier in a memory.

4. (Currently Amended) The method of claim 3, wherein the previously-labeled set of electronic communications includes electronic communications previously labeled by the user.

5. (Currently Amended) The method of claim 4, wherein the previously-labeled set of electronic communications includes electronic communications suspected by the system to be possibly mis-labeled by the user.

6. (Currently Amended) The method of claim 3, wherein the previously-labeled set of electronic communications includes electronic communications previously labeled by the system.

7. (Currently Amended) The method of claim 3, wherein the previously-labeled set of electronic communications includes electronic communications previously labeled by a user and electronic communications previously labeled by the system.

8. (Currently Amended) The method of claim 3 wherein presenting the electronic communications to the user includes :

assessing a value that labeling a set of electronic communications from each group will provide to the classifier being developed; and

selecting a next group for labeling based upon the greatest respective value that will be provided to the classifier being developed from the assessing .

9. (Currently Amended) The method of claim 3 wherein presenting the electronic communications to the user includes :

assessing a value that labeling a set of electronic communications from each group will provide to the classifier being developed; and

selecting a next group for labeling based upon achieving known performance bounds for the classifier.

10. (Previously Presented) The method of claim 3 further comprising developing an expression of labeling criteria in an interactive session with the user.

11. (Previously Presented) The method of claim 10, wherein the interactive session includes posing hypothetical questions to the user regarding what type of information the user would consider relevant.

12. (Original) The method of claim 11, wherein the hypothetical questions elicit "yes", "no" and "unsure" responses from the user.

13. (Original) The method of claim 11 wherein subsequent questions are based, at least in part, upon the answers given to previous questions.

14. (Previously Presented) The method of claim 11 wherein developing an expression of labeling criteria produces a criteria document.

15. (Currently Amended) The method of claim 14 wherein the expression and/or the criteria document include a group of keywords and/or phrases for use by the system in automatically labeling electronic communications.

16. (Previously Presented) The method of claim 10 wherein developing an expression of labeling criteria produces a criteria document.

17. (Previously Presented) The method of claim 16 wherein the criteria document includes a list of items that are considered relevant and a list of items that are considered irrelevant.

18. (Currently Amended) The method of claim 17, wherein presenting the electronic communications to the user includes querying the user to identify which item(s) influenced the label on a user-labeled electronic communication.

19. (Currently Amended) The method of claim 16, wherein ~~the~~ at least one of the expression ~~and/or~~ the criteria document include at least one of a group of keywords ~~and/or~~ phrases for use by the system in automatically labeling electronic communications.

20. (Previously Presented) The method of claim 10 wherein the interactive session is conducted prior to presenting the electronic communications to the user.

21. (Currently Amended) A computer ~~assisted~~/implemented method for developing a classifier for classifying electronic communications comprising:

(a) developing an expression of labeling criteria in an interactive session with a user, wherein the interactive session includes querying a user to identify an irrelevancy keyword indicative of an irrelevant electronic communication and receiving a user identification of the irrelevancy keyword;

(b) presenting electronic communications to the user for labeling as relevant or irrelevant, wherein the electronic communications are user-generated;

(c) developing a classifier for classifying electronic communications based upon the irrelevancy keyword and the ~~relevant/irrelevant~~relevant labels and the irrelevant labels assigned by the user; ~~and~~

(d) deploying the classifier for use in classifying electronic communications based upon the irrelevancy keyword and the ~~relevant/irrelevant~~relevant labels and the irrelevant labels;

(e) storing a set of electronic communications labeled by the classifier in a memory;
and

wherein at least one of (b) and (c) use the expression of labeling criteria developed in ~~the~~ (a).

22. (Previously Presented) The method of claim 21, wherein the interactive session includes posing questions to the user regarding what type of information the user would consider relevant.

23. (Original) The method of claim 22, wherein the questions elicit "yes", "no" and "unsure" responses from the user.

24. (Previously Presented) The method of claim 22 wherein subsequent questions are based, at least in part, upon the answers given to previous questions.

25. (Previously Presented) The method of claim 22 wherein the questions are structured from several dimensional levels of relevance, including a first dimension of question segments on a topic, a second dimension of question segments on an aspect of the topic and a third dimension of question segments on a type of discussion.

26. (Currently Amended) The method of claim 25, wherein:

the first dimension of question segments on a topic include one or more of the following segments: a first segment concerning a client's product and a second segment concerning a client's competitors;

the second dimension of question segments on a topic include one or more of the following segments: a third segment concerning a feature of the ~~first segment~~topic, a fourth segment concerning the ~~first segment~~topic itself, a fifth segment concerning corporate activity of the ~~first segment~~topic, a sixth segment concerning price of the ~~first segment~~topic, a seventh segment concerning news of the ~~first segment~~topic and ~~an eighth~~an eighth segment concerning advertising of the ~~first segment~~topic; and

the third dimension of question segments on a topic include one or more of the following segments: a ninth segment concerning a mention of the second dimension segment, a tenth segment concerning a description of the second dimension segment, ~~an eleventh~~an eleventh segment concerning a usage statement about the second dimension segment, a twelfth

segment concerning a brand comparison involving the second dimension ~~segment of~~
questions segments, and a thirteenth segment concerning an opinion about the second
dimension segment.

27. (Previously Presented) The method of claim 21 wherein developing the
expression of labeling criteria produces a criteria document.

28. (Original) The method of claim 27 wherein the criteria document includes
a list of items that are considered relevant and a list of items that are considered irrelevant.

29. (Currently Amended) The method of claim 28 wherein the criteria document
includes a group of keywords for use by the system in automatically labeling electronic
communications.

30. (Previously Presented) The method of claim 28, wherein presenting the
electronic communications to the user includes querying the user which items influenced the
label on a user-labeled communication.

31. (Currently Amended) The method of claim 21 wherein the expression of
labeling criteria includes a group of keywords and/or phrases for use by the system in
automatically labeling electronic communications.

32. (Currently Amended) The method of claim 31 wherein the group of keywords
is also for use by the system in gathering electronic communications.

33. (Currently Amended) A computer ~~assisted~~/implemented method for developing a classifier for classifying electronic communications comprising:

- (a) defining a domain of electronic communications on which a classifier is to operate, wherein the electronic communications are user-generated;
- (b) collecting a set of electronic communications from the domain;
- (c) eliciting labeling ~~communication~~ criteria from a user by querying a user to identify a keyword indicative of an irrelevant electronic communication and receiving a user identification of the keyword;
- (d) labeling, by the system, electronic communications from the set of electronic communications according, at least in part, to the labeling ~~communication~~ criteria elicited from the user;
- (e) labeling, by the user, electronic communications from the set of electronic communications;
- (f) building the electronic communications classifier according to a combination of labels applied to electronic communications in (d) and (e); ~~and~~
- (g) deploying the classifier for use in classifying electronic communications based upon the combination of labels; and
- (h) storing a labeled set of electronic communications labeled by the classifier in a memory.

34. (Currently Amended) The computer implemented method of claim 33, wherein (d) and (e), and (f) includes selecting electronic communications for labeling by the

user targeted to build the electronic communications classifier within known performance bounds.

35. (Currently Amended) The computer implemented method of claim 34, wherein selecting electronic communications for labeling by the user selects electronic communications from groups of electronic communications including:

a training set group of electronic communications, the training set group of electronic communications being selected by an active learning algorithm;

a test set group of electronic communications for testing the accuracy of a current state of the classifier; and

a previously-labeled set of electronic communications previously labeled by at least one of the user, the system and another user.

36. (Currently Amended) The computer implemented method of claim 34, wherein selecting electronic communications for labeling by the user selects electronic communications from groups of electronic communications including:

a training set group of electronic communications selected by an active learning algorithm;

a system-labeled set of electronic communications previously labeled by the system;

a test set group of electronic communications for testing the accuracy of a current state of the classifier being developed;

a faulty set of electronic communications suspected to be previously mis-labeled by the user; and

a random set of electronic communications previously labeled by the user.

37. (Currently Amended) The computer implemented method of claim 33, wherein the ~~communication~~labeling criteria elicited in the eliciting of (c) is used, in part, to determine electronic communications to collect in the collecting of (b).

38. (Previously Presented) The computer implemented method of claim 37, wherein the eliciting (c) involves an interactive session with the user.

39. (Currently Amended) The computer implemented method of claim 37, wherein the ~~communication~~labeling criteria elicited in the eliciting (c) is used, in part, by the system to label electronic communications in the labeling (d).

40. (Previously Presented) The computer implemented method of claim 39, wherein the eliciting (c) involves an interactive session with the user.

41. (Previously Presented) The method of claim 33, wherein the building (f) involves an active learning process.
42. (Currently Amended) The computer implemented method of claim 33, wherein the ~~communication~~labeling criteria elicited in the eliciting (c) is used, in part, by the system to label electronic communications in the labeling (d).
43. (Previously Presented) The computer implemented method of claim 33, wherein the eliciting (c) involves an interactive session with the user.
44. (Previously Presented) The method of claim 43, wherein the interactive session includes posing questions to the user regarding what type of information the user would consider relevant.
45. (Original) The method of claim 44, wherein the interactive session also allows the user to provide keywords based upon a criteria the user considers relevant.
46. (Canceled)
47. (Original) The method of claim 44, wherein the questions elicit "yes", "no" and "unsure" responses from the user.
48. (Previously Presented) The method of claim 43, wherein the building (f) involves an active learning process.

49. (Currently Amended) A computer memory containing a software program including instructions for implementing a method for developing a classifier for classifying electronic communications comprising:

querying a user to identify a keyword indicative of an irrelevant electronic communication;

receiving a user identification of the keyword;

(a) presenting electronic communications to a user for labeling as relevant or irrelevant, the electronic communications being selected from groups of user-generated electronic communications including:

a training set group of electronic communications selected by an active learning algorithm;

a test set group of electronic communications for testing the accuracy of a current state of a classifier being developed; and

a previously-labeled set of electronic communications previously labeled by at least one of the user, the system and another user;

(b) developing the classifier for classifying electronic communications based upon the keyword and the ~~relevant/irrelevant~~relevant labels and the irrelevant labels assigned by the user during presenting electronic communications to the user; ~~and~~

(c) deploying the classifier for use in classifying electronic communications based upon the ~~relevant/irrelevant~~relevant labels and the irrelevant labels; and

storing a set of electronic communications labeled by the classifier in a memory.

50. (Currently Amended) A computer memory containing a software program including instructions for implementing a method for developing a classifier for classifying electronic communications comprising:

(a) developing an expression of labeling criteria in an interactive session with the user, wherein the interactive session includes querying a user for an irrelevancy keyword indicative of an irrelevant electronic communication and receiving a user identification of the irrelevancy keyword;

(b) presenting electronic communications to a user for labeling as relevant or irrelevant, wherein the electronic communications are user-generated; and

(c) developing a classifier for classifying electronic communications based upon the irrelevancy keyword and the ~~relevant/irrelevant~~relevant labels and the irrelevant labels assigned by the user; ~~and~~

(d) deploying the classifier for use in classifying electronic communications based upon the irrelevancy keyword and the ~~relevant/irrelevant~~relevant labels and the irrelevant labels; and

storing a set of electronic communications labeled by the classifier in a memory;

wherein at least one of (b) and (c) use the expression of labeling criteria developed in (a).

51. (Currently Amended) A computer system memory containing a software program including instructions for implementing a method for developing a classifier for classifying electronic communications comprising:

(a) defining a domain of electronic communications on which a classifier is to operate, wherein the electronic communications are user-generated;

(b) collecting a set of electronic communications from the domain;

(c) eliciting labeling ~~communication~~ criteria from a user by querying a user for an irrelevancy keyword indicative of an irrelevant electronic communication and receiving a user identification of the irrelevancy keyword;

(d) labeling, by the computer system, electronic communications from the set of communications according, at least in part, to the labeling ~~communication~~ criteria elicited from the user;

(e) labeling, by the user, electronic communications from the set of electronic communications;

(f) building the electronic communications classifier according to a combination of labels applied to electronic communications in (d) and (e); ~~and~~

(g) deploying the classifier for use in classifying electronic communications based upon the combination of labels; and

storing a set of electronic communications labeled by the classifier in a memory.